

**Social Security Earnings Test and the Labor Supply of the Elderly:
New Evidence from Unique Survey Responses in Japan***

by

Satoshi Shimizutani

This version: May 2009

* This project was undertaken in a research project exploring the interaction between the pension program, retirement, and the well-being of the elderly supported by Health and Labour Sciences Research Grants and Project on International Equity. We are grateful to the Ministry of Health, Labour and Welfare for providing us a micro-level dataset. We also thank Hidehiko Ichimura, Takashi Oshio, Izumi Yokoyama for their very constructive comments, which improved the manuscript substantially. The views expressed in this paper are those of the author and do not represent any institution.

⁺ Corresponding author: Senior Research Fellow, Institute for International Policy Studies; Toranomom 30 Mori Building 6th floor, Toranomom 3-2-2, Minato-ku, Tokyo Japan 105-0001; Tel: +81-3-5404-6651, FAX: +81-3-5405-6650; E-mail: sshimizutani@iips.org.

Abstract

Although there exists a large volume of literature on the subject, a consensus on the labor supply effects of the social security earnings test for the elderly has yet to be reached. This study proposes an alternative approach of utilizing direct responses to a survey on the earnings test, a unique feature of our dataset compiled by the Japanese Government, to provide new evidence on the sensitivity of the labor supply decision of workers aged between 60 and 64 with respect to the earnings test. Our empirical results show that a large proportion of these workers are discouraged from working or reduce their working hours, even after correcting for observed attributes of individuals who reported either affected or unaffected. In addition, the revision of the test rules in 1995 did not alter the labor supply of the elderly.

Keywords: social security earnings test, labor supply of the elderly, Japan, wage distribution, DiNardo-Fortin-Lemieux decomposition.

JEL Classification Codes: H55, J26.

1. Introduction

Despite the tremendous volume of published literature on the subject, it is fair to state that there is no consensus on the effect of the social security earnings test on the labor supply of elderly workers. The standard textbook theory on labor supply explains that the earnings test kinks the budget constraint for an individual at a threshold and that a change in the test rule shifts the budget constraint. Since the total effect depends on the relative size of income and substitution effects (Borjas (2005)), to identify the labor supply effect of the earnings test on a net basis, an empirical investigation is required. While a majority of the existing literature found a small effect,¹ recent studies revealed a large discouraging effect of the earnings test on labor supply (Friedberg (2000), Haider and Loughran (2008)), which makes empirical findings rather mixed. The inconclusiveness is troublesome especially for countries that experience a rapidly aging society and find stimulus of the elderly labor supply to be an urgent policy issue.

Japan is not an exception. One popular view in domestic policy debates is that the social security earnings test discourages paid work for the elderly through a high effective tax rate on income. When the sum of labor income and public pension

¹ Burtless and Moffitt (1985) and Gustman and Steinmeier (1985) are earlier works that conclude that the elimination of the earning tests would have only minor effects on the labor supply.

income exceeds a certain exemption threshold, the earnings test reduces immediate payments of pension benefits to the beneficiaries of the Employee's Pension Insurance (EPI; *Kosei Nenkin*) program, which covers about half the pensioners in Japan. The earnings test is often viewed as "punishment" on the elderly labor supply.

There are two strands of studies to examine the labor supply effects (see Gruber and Orszag (2003) for the literature). One is to employ a "bunch analysis" which examines the concentration of the earnings distribution below the threshold of the earnings tests. An example is Friedberg (2000) who confirmed that the bunching shifted in response to the revision of the earnings test in the US. The other type of research employs sophisticated econometric models to explore the aggregate impacts of the kinked budget constraint in the earnings test on the conditional hours worked by older workers. Again, Friedberg (2000) showed that the structural estimation yielded sizable impacts of removal of the earnings test on workers aged 65 and over.

Those two approaches are surely effective but not perfect. One caveat of the bunch analysis is that it is hard to detect the labor supply effect when the response to the earnings test in survey data is obfuscated by measurement errors and labor market rigidities. Indeed, Haider and Loughran (2008) discussed this point and found a consistent and substantial response to the earnings test, especially for younger men, if

those elements are taken into account. On the other hand, structural estimations require a variety of assumptions of fundamental parameters which are not necessarily based on solid empirical research and but are likely to influence the magnitude of the labor supply effect.

These weaknesses in econometric analyses in the previous studies are partly responsible for the inconclusiveness regarding the effect of the social security earnings test on the labor supply.² In this study, we propose an alternative approach i.e., a direct response to a survey on the effect of the earnings test, unique in the literature, to detect the labor supply decision of the elderly subject to the earnings test. The *Survey on Employment of the Elderly* (Konenreisha Shugyo Jittai Chosa, henceforth SEE) compiled by the Japanese government is a large scale nationwide survey that contains rich information both on employment status and eligibility for social security programs. The SEE explicitly asks the respondents how the earnings test discouraged them from working. This feature enables us to identify a discouraging effect for those who quit working and the degree to which they circumvent a higher effective tax on labor

² The number of studies on this topic is tremendous. Recent studies pertaining to the US are represented by Friedberg (2000), Haider and Loughran (2008), and Gruber and Orszag (2003), the last of which provides a very concise literature survey. Outside the US, Disney and Tanner (2002) examined the effect of the earnings test removal in the UK and Baker and Benjamin (1999) in Canada. In Japan, while a majority of studies found the significant labor supply effect of the earnings test including Abe (1998), Ogawa (1998) and Iwamoto (2000) found a sizable effect on labor supply.

income for incumbent workers and distill the labor supply effect of the earnings test for workers aged between 60 and 64.

The advantages to examining direct survey responses to the earnings test are summarized in the three aspects by which we aim to contribute to the literature in an innovative way. First, this approach is exempt from misidentification when labor market rigidity exists. A bunch analysis implicitly assumes that a labor market is flexible and that workers are able to adjust their earnings and working hours. However, in reality, some workers may not be able to adjust their working hours to earn just below the threshold of the earnings test due to inflexibility in the work schedule. If this is the case, the earning concentration below the threshold is obscured and the discouraging effect is underestimated. This issue is emphasized in a recent work by Haider and Loughran (2008). The direct response approach enables us to identify those workers who want to work more hours but earn far below the threshold under earnings test, which is disregarded in a bunch analysis.

Second, we are able to directly measure the discouraging effect for workers who gave up working entirely, the people who are most seriously affected by the earnings test. In most cases, a bunch analysis examines wage concentration below the threshold of the earnings test for incumbent workers only and implicitly completely disregards

discouraged workers, which again understates the labor supply effect. In other words, our approach reveals the total effect of the earnings test on the labor supply of the elderly.

Third, the direct survey responses enable us to distill the labor supply effect of the earnings test separately from the remaining factors. In most cases, it is difficult to identify whether lower earnings are caused by the earning test or other factors and we are forced to infer the effect by controlling for factors responsible for earnings using unobserved information, which might be insufficient to derive the effect of interest. The direct survey responses to the earnings test facilitate discerning the effect of the earnings test on labor supply.

While we see some notable merits in the use of direct responses to the earnings test, our approach is not claimed to be the best way to evaluate the effect of the earnings test on labor supply. Indeed, we acknowledge some disadvantages. One possible drawback is that the survey response is likely to be subjectively biased. Another possible disadvantage is that the survey responses are also likely to be contaminated by reporting errors. In order to address these issues, we examine whether the gap in the distribution of wage plus pension benefits, the object of the earnings test, is observed between those reporting the discouraging effect and those not reporting,

even after controlling for observed characteristics of individuals. We examine the difference in wage distributions using the methodology of DiNardo Fortin and Lemieux (1996), which enables us to construct a counterfactual distribution as if the observable attributes of both groups were homogenous.³ Nevertheless, we propose that use of the direct survey responses to the earnings test complements traditional approaches.

This paper proceeds as follows. The next section briefly describes the earnings rule for workers aged 60–64 in Japan and Section 3 explains the dataset used in this study. Section 4, the main body of the paper, presents the results of the direct survey responses in the 2000 survey and compares wage plus pension benefits distributions of those reporting and not reporting discouraging effects of the earning tests. Section 5 also examines the effect of the change in the revision rule in 1995 and contrasts the results with those in the previous section. The last section concludes.

2. Social security earnings test for workers aged 60–64 in Japan

The social security earnings test, which is called the *Zaishoku Rorei Nenkin* program, was first introduced in Japan in 1965 and has been revised every four to six

³ Lemieux (2002) applied this methodology to examine the effect of change in minimum wage on wage distribution.

years. The format is different for workers aged between 60 and 64 and those aged 65 and over. Since this study takes advantage of micro-level data from the 1996 and 2000 SEE surveys, we will briefly describe the earnings test format for workers aged between 60 and 64 since the 1990s. We should remember that the earnings test rule is applied only to the beneficiaries of the Employees' Pension Insurance (EPI: *Kosei Nenkin*) program which is applied to employees in the private sector, but not to the other public pension programs.⁴ The object of the earnings test includes first-tier pension benefit for workers aged 60–64 while this is not the case for those aged 65–69.

In this study, we confine our interest to EPI pensioners aged 60–64, whose labor supply is presumably most seriously affected by the earnings test. The earnings test for those workers has multiple thresholds and the reduction rates differ across earnings brackets. In 1989, the object of the earnings test was labor income.⁵ In 1995, the object of the earnings test became the sum of labor income and public pension benefits (first-tier and second-tier). At the same time, the effective rate was simplified to 20 percent for less than 220,000 yen, 60 percent for 220,000 yen–340,000 yen and 100

⁴ National Pension Insurance (NPI: *Kokumin Nenkin*) for self-employed persons and the Mutual Aid Insurance (*Kyosai Nenkin*) for employees in the public sector and private schools.

⁵ The effective rate for workers aged 60–64 was revised as follows: 20 percent for less than 95,000 yen, 30 percent for 95,000 yen–114,000 yen, 40 percent for 114,000 yen–138,000 yen, 50 percent for 138,000 yen–165,000 yen, 60 percent for 165,000 yen–185,000 yen, 70 percent for 185,000 yen–210,000 yen, 80 percent for 210,000 yen–230,000 yen, and 100 percent for 230,000 yen and over.

percent for 340,000 yen and over. We will examine the effect of the 1995 reform on elderly labor supply to complement our analysis in Section 5.⁶

3. Data description

This study uses micro-level data from the SEE compiled by the Ministry of Health, Labour and Welfare of the Japanese government. The SEE is performed every four or five years. In this study, we use micro-level data from the SEE surveys compiled in 1996 and 2000 which contain direct responses regarding the earnings test. The individuals in the sample were aged between 55 and 69 and randomly chosen from all regions in Japan. The total sample size was 21,219 in the 1996 survey and 19,595 in 2000. The number of workers aged 60–64 in the sample is 7,383 in 1996 and 6,692 in 2000. To our knowledge, the SEE is only large-scale dataset compiled by the Japanese government and contains a selection of variables that are indispensable to examination of the labor supply effect of the earnings test, including the individual's demographics

⁶ Although beyond the scope of this study, the earnings test rule for those aged 60-64 was revised again in 2002 and the threshold 60 percent rate increased from 340,000 yen to 370,000 yen. The rule was slightly revised in 2004 to include bonuses as labor income and change the reduction rate correspondingly, but there has been virtually no change in the rule since 2002. The earnings were not tested for workers aged 65 and over after the elimination of the earnings test in 1985. In 2002, the test revived and the workers who earned more than 370,000 yen from labor income and second-tier pension benefit face a marginal tax rate of 50 percent (first-tier pension benefit is not tested for those workers aged 65 and over). In 2007, the earnings test started to be applied to workers aged 70 and over. For those workers aged 65 and over, the first-tier benefit is not tested. See Shimizutani and Oshio (2008) for analysis for the workers aged 65 and over.

and employment status (age, sex, health status, type of job, monthly wage, working days per week, and hours worked per day, and others) as well as social security eligibility and benefits.⁷

Another feature of the SEE is that it asks the respondents who are beneficiaries of the EPI program explicitly whether the social security earnings test discourages them from working. Since the survey restricts the respondents of the question to those who are eligible for the EPI program, there is no risk of misidentifying other respondents as EPI beneficiaries.⁸ The 1996 survey asked the effect of the *change* in the earnings test rule, i.e., the modified earnings test, in the 1995 reform on labor supply, while the 2000 survey inquires regarding the effect of the current earnings test *per se* which was effective in 2000 on labor supply decision. We confine our analysis to the respondents of these questions and use the respondents who are eligible for Mutual Aid Insurance (MAI) exempt from the earnings test to complement the analysis. Details of the questions will be provided in later sections.

In what follows, we examine the effect of the earnings test rule on labor supply

⁷ There are other large scale data sets on employment collected by the Japanese government, represented by *Labor Force Survey* (Rodo Ryoku Chosa) or *Basic Survey on Employment Structure* (Shugyo Kozo Kihon Chosa) but there is no information on pension eligibility or the direct survey response which is indispensable to this study.

⁸⁸ Some people who are eligible for the EPI program decline their benefit when they are entitled to it and instead postpone. In this case, we are not able to identify the persons who are eligible using the information on actual receipt of the benefit. As a result, we implicitly exclude those who are eligible but do not currently receive any EPI benefits.

for workers aged 60–64 using the 2000 survey in Section 4 and then turn to examine the effect of the 1995 reform using the 1996 survey in Section 5.

4. Direct survey response and distribution of wage and pension benefits in 2000

In this section, we utilize the respondents' direct survey responses to the earnings test in the 2000 survey. This survey asked the EPI beneficiaries as follows: "Do you restrict working hours or days due to a reduction in or no receipt of EPI benefits under the social security earnings test?" Each respondent was asked to choose one of three answers: (1) I did not work at all because of the earnings test, (2) I restricted working hours or days because of the earnings test, or (3) I did not adjust employment in consideration of the earnings test. The number of respondents eligible for the EPI pension program is 3,555, of whom 2,022 answered these questions in the 2000 survey. Further, we confine our sample to the individuals whose wage and nonwage incomes are available in order to be able to compute the wage distribution.

Table 1 reports the results on the labor supply response by three groups; "no work group" choosing answer (1), "restricted group" choosing answer (2), and "no effect group" choosing answer (3). In the 2000 survey, the SEE asked respondents to report their monthly labor income (in a 10,000 yen unit integer) earned in September,

just one month before the October survey was conducted.⁹ It is intuitive to predict that wage income and labor supply is larger for the “no effect group” and smaller for the “no work group,” and indeed the prediction is confirmed by data. The average monthly wage income is the smallest at only 14,000 yen for the “no work group,” and the largest at 272,200 yen for the “no effect group,” while the “restricted group” is in between at 135,800 yen. While we see a different pattern in nonwage income (the average of nonwage income is 150,000–160,000 yen for the “no work group” and the “restricted group” while it is low at 8,400 yen for the “no effect group”), the sum of wage income and first- and second-tier public pension income, the object of the earnings test, differs substantially across the groups: 115,900 yen for the “no work group,” 259,500 yen for the “restricted group,” and 342,300 yen for the “no effect group.” This order is similarly and clearly observed in the labor supply. The average working hours per day is less than 0.5 hours for the “no work group” but 6.7 hours for the “restricted group” and 7.5 hours for the “no effect group,” implying that the workers in the last category work on a full-time basis. This is also the case for working days per week. The average working days per week are only 0.3 days for the “no work group” but 4.3 and 5.2 days, respectively, for the remaining groups. As a result, we see

⁹ We believe that the close proximity in time renders the information reliable.

a substantial difference in monthly working hours: 2.1 hours, 28.8 hours, and 39.9 hours for the three groups.

These observations show that the earnings test strongly discourages labor supply in a subgroup of the elderly. The sum of wage and pension income of the “no work group” is one third that of the “no effect group” and that for the “restricted group” is two thirds. This difference is also confirmed by comparing the figures of the “no effect group” with those of MAI beneficiaries who are exempt from the earning test. The table shows that the sum of wage and pension benefits is comparable between the “no effect group” and MAI beneficiaries, confirming that the wage and pension benefits are substantially smaller for the “no work group” or “restricted group.” Moreover, what we should notice is that the average of the sum of wage and pension income is far below the threshold of the earnings test. This average of the value for the “restricted group” or the “no work group” is beyond the lower threshold (220,000 yen).

However, we need to examine not only the average but also the distribution to capture the effect of the earnings test on labor supply. In what follows, we reorganize the three groups into two as follows: the “affected group” consisting of the “no work group” and “restricted group,” and the “unaffected group” which stands for the “no effect group.” The share of each group in the pooled sample is close to 50 percent (898

versus 892). Figure 1 illustrates the distribution of the sum of monthly wage income and first- and second-tier pension income. The range in the X axis is from zero to 0.8 million yen and each cell stands for each 10,000 yen bracket.¹⁰ If the social security earnings tests restrain labor supply, we would observe a bunch below the thresholds (220,000 yen and 340,000 yen) not for the “unaffected group” but for the “affected group.” Note that the first- and second-tier benefits are computed as full benefits without the earnings test before they were reduced, which is available in the 2000 SEE survey, in order to examine behavioral responses of labor supply to the earnings test.

We observe clearly that the whole distribution is located more to the left hand side (close to zero) for the “affected group,” implying that the wage plus pension income is lower for that group, which is consistent with the discussion on the average reported in Table 1. A closer look shows no bunch either below 220,000 yen (the share of 190,000, 200,000 and 210,000 yen is 8.7 percent in total and that of 220,000, 230,000 and 240,000 is 9.6 percent) or below 340,000 yen (the share of 310,000, 320,000 and 330,000 is 3.0 percent and that of 340,000, 350,000 and 360,000 is 3.23 percent) for the “affected group.” Figure 2 depicts the reduction rate which is defined

¹⁰ We set the upper limit at 0.8 million yen to make the scale of the figures the same. The proportions of individuals whose sum of wage and pension benefits is greater than 0.8 million and thus who do not appear in the figures are 0.032 percent for the “affected group” and 4.37 percent for the “unaffected group.”

as the share of the amount reduced under the earnings test out of the pension benefits without the test. We see a striking difference between the two groups. For the “affected group,” the most frequent is a 20 percent reduction and, in most cases, the reduction rate is less than 50 percent. In contrast, more than 10 percent of the “unaffected” group has a reduction rate of 100 percent, followed by 50 percent reduction.

These findings provide two important observations. First, figures of the workers who reported being undiscouraged by the earnings test indeed support their response that they are not concerned about the threshold of the earnings test. They earn above the threshold and the reduction rate is high. We do not see any bunches on the distribution of the sum of wage and pension income and a large portion of the eligible pension income is reduced. Second, the wage and pension income and working hours are much lower for the affected group, a very important observation relating to stimulus of the elderly labor supply. What is important is that the sum of wage and pension income does not form a bunch below the thresholds for the “affected group,” implying that workers who express concern about the earnings test are not able to adjust their wages close to the thresholds for certain reasons such as rigid working hours. In such a case, a bunch analysis has the pitfall of overlooking the effect of the earnings test.

However, it is too early to conclude that the social security earnings test strongly discourages half of the individuals aged between 60 and 64 since those people are inherently more likely to respond that their labor supply decision is discouraged by the earnings test. If this is the case, the discouraging effect of the earnings test would be overestimated. Thus, we need to explore whether the difference in the distribution of wage plus pension benefits is observed between those reporting the discouraging effect and those not reporting it, even after controlling for observed characteristics of workers.

Indeed, Table 2 demonstrates that the two groups are not homogenous in terms of the direct response to the earnings test and have different characteristics. The proportion of males is larger for the “unaffected group.” With the dominant age cohort at 60, the respondents’ age is younger for the “unaffected group” while it is 63 for the “affected group.” The “unaffected group” has a higher educational attainment as clearly seen in the share of university graduates and above: 9 percent versus 17 percent for each group. The difference is also observed in the self-report health status, which is worse for the “affected group.” The share of those who responded they are healthy is 71 percent for the “affected group” and 86 percent for the “unaffected group” and that of those who responded they are sick is 10 percent and 2 percent, respectively. A

similar pattern is observed for the subjective possibility in terms of physical status to work: the share of those who are physically able to work is only 32 percent for the “affected group” but 67 percent for the “unaffected group.” One half of the individuals in the “affected group” responded they are able to work but that work conditions were a factor. Finally, we do not observe any difference in family size.

Thus, we should explore the possibility that the difference in the distribution of wage plus pension income is still observed even if the observable characteristics were homogenous between the groups. Concretely, we examine the difference in wage plus pension income distribution using the methodology of DiNardo Fortin and Lemieux (1996), which enables us to construct a counterfactual distribution as if the observable attributes of both groups were homogenous. This is a semi-parametric approach and visually decomposes the change in wage distribution into two parts: the change in the distribution of the attributes and the change in the effect of attributes on distribution of wage and pension income.

Concretely, the decomposition approach is described as follows. Denoting i as a group, the distributions of wage and pension income for the “affected group” and the “unaffected group” are written as

$$f^{affected}(Y) = \int f^{affected}(Y|X)h(X|i=affected)dX,$$

$$f^{not_affected}(Y) = \int f^{not_affected}(Y|X)h(X|i=not_affected)dX,$$

respectively, where $f^{affected}(Y|X)$ is the mechanism to determine the wage and pension income of the “affected group” that maps workers’ characteristics X to the wage distribution Y and $f^{not_affected}(Y|X)$ is that of the “unaffected group.” Moreover, what the distribution of the sum of wage and pension income would be if the distribution of X of the “affected group” is that of the “unaffected group” is written as

$$f_{affected}^{not_affected}(Y) = \int f^{not_affected}(Y|X)h(X|i=affected)dX.$$

The DiNardo Fortin and Lemieux approach employs a “re-weighting” method to estimate the counterfactual distribution. The counterfactual distribution can be rewritten as:

$$\begin{aligned} f_{affected}^{non_affected}(Y) &= \int f^{non_affected}(Y|X)h(X|i=affected)dX \\ &= \int \omega f^{non_affected}(Y|X)h(X|i=non_affected)dX, \end{aligned}$$

where $\omega \equiv \frac{h(X | i = affected)}{h(X | i = non_affected)}$. The Bayesian rule produces

$$\omega = \frac{P(i = affected | X)}{P(i = non_affected | X)} \frac{P(i = non_affected)}{P(i = affected)}$$

where the conditional probabilities, $P(i = affected | X)$ and $P(i = non_affected | X)$ are propensity scores for the specific observations for those who are affected and those who are unaffected, respectively, conditioned on X , which are calculated by the logit model in this analysis. The terms $P(i = affected)$ and $P(i = non_affected)$ are calculated based on the proportion of the observations for each group. The counterfactual distribution is computed by the kernel density estimation, using calculated weight ω . The kernel density is also useful to adjust for reporting errors in this study. To make the results comparable with those from the histogram analysis, we take the level of wage and pension income as the dependent variable but the results are unchanged when we take the logarithm of the sum as the dependent variable.

In the analysis, we compare the actual distribution for the “affected group” with the counterfactual distribution which is defined as what the density of wage would have been if the attributes of workers of the “unaffected group” were those of the “affected group.” If the difference in the observable attributes between the two groups produces a difference in the distributions, the counterfactual distribution will overlap with the distribution of the “affected group.” If this is not the case, however, and the

counterfactual distribution still remains to the right of that of the “affected group,” the discouraging effect of the earnings test between the two groups would not be explained by the difference in the characteristics of the two groups, which is observable.

Figure 3 reports the actual distributions of the two groups and the counterfactual distribution. When comparing the actual distributions in the two groups, we see that the “peak” in the distribution of the affected group is located to the left of that of the unaffected group, which conforms to the histogram analysis in Figure 1. What is interesting is the comparison between the actual distribution of the affected group and the counterfactual distribution, both of which are based on the same observable attributes of the affected group. We still see that the distribution of the counterfactual distribution is located to right of that of the “affected group.” This finding shows that, even after correcting the selection in the observable attributes between the two groups, the sum of wage and pension income is smaller for the affected group and the earnings test discourages a subgroup of the individuals in the sample from working more. From a closer look, we observe that the peak of the affected group is located near 220,000 yen, the lower threshold of the earnings test. Moreover, the cumulative density is larger for the unaffected group above 340,000 yen, the upper threshold. These observations reinforce the finding on the discouraging effect of the earnings test on labor supply for

the individuals reporting that they were discouraged by the test.

In the analysis mentioned above, we included the “no work” group in the sample. In order to explore the possibility that the finding is altered by those who do not work at all, we exclude those workers and perform the same decomposition analysis for the “restricted group” and the “unaffected group.” Figure 4 describes the actual distributions of the sum of wage and pension income of those two groups and the counterfactual distribution as if the observed attributes of the individuals in the “unaffected group” were the same as those of the “restricted group.” In the figures, we see that the counterfactual distribution overlaps that of the unaffected group. This implies that the difference in the distributions of the two groups does not come from the difference in the characteristics of the individuals in the sample. In other words, even if the observable characteristics are homogenous between the two groups, the distribution of the unaffected group is unchanged and the gap between the two groups remains.

Together with the findings in Figure 3, we confirm that the discouraging effect of the earnings test is clearly observed for the affected group and the difference in the distribution of the affected and unaffected groups does not stem from different characteristics of the individuals in each group. These observations show that, even in

the case that we do not find a bunch below a threshold of the earnings test, we still find a large discouraging effect of the earnings test, which has been largely ignored by the previous work on the topic.

5. Direct survey response to the 1995 reform in the 1996 survey

In this section, we take advantage of the direct response to the 1995 reform addressed in the 1996 survey in order to complement the findings in the 2000 survey. Some aspects of the earnings test for those aged 60 and 64 were revised in 1995. First, the object of the test became the sum of labor income and first-tier and second-tier pension benefits, not just the amount of labor income. Second, more importantly, the reduction rate (only three rates after the reform) and the thresholds (two after the reform) were greatly simplified to erase the ranges of the income bracket for which, after reduction of the pension income, the net income decreased. Before the reform, if a worker earned more, his net income declined since the effective reduction rate of the earnings test was high.

The 1996 survey asked “Did the reform in 1995 affect your employment status?” Each respondent was asked to choose one of four answers: (1) yes, I began to work, (2) yes, I now earn more than before, (3) yes, I moved to a firm which provides a higher

wage, (4) no, the reform did not affect my employment at all. The difference between the 1996 survey and the 2000 survey is that the 1996 survey asked the respondents about their response to the change in the earnings test rule while the 2000 survey asked about their response to the incumbent earnings test rule *per se*. Since the question was asked only of respondents who actually received a reduced amount of public pension income under the earnings test, the number of the respondents who answered was 262 in 1996.

In what follows, we organize two groups: the “positively affected group” consisting of the respondents who chose (1), (2), or (3) and the “no effect group” for those who chose (4). Table 3 reports the summary statistics. First, we observe that the share of those who were encouraged to work more due to the 1995 reform is small at 13 percent while the remaining did not alter their labor supply. Second, in terms of wage income, nonwage income, and pension benefits, the individuals in each group are comparable. Third, the proportions of males and the younger age bracket are larger for the positively affected group, implying those workers are more willing to work more. Fourth, we do not see a large difference in health status but the share of those who self-report their physical status as able to work on a full-time basis is larger for the positively affected group.

These observations show that the 1995 reform stimulated the labor supply of a small proportion of the workers who are male, younger (in their 60s), and physically able to work on a full-time basis, but the effect is limited. One possible explanation is that workers are not able to adjust their labor supply freely. If they were able to do so, we would see a bunch below a threshold of the earnings test. However, we found that this is not the case in reality, implying that small changes of the test rule do not alter labor supply behavior. Together with our findings using the 2000 survey, the findings of the 1996 survey indicate that the discouraging effect of the social security earnings test is large and that partial reform had only a limited effect.

6. Concluding remarks

The large volume of literature has not reached a consensus on the labor supply effect of the social security earnings test for the elderly. We propose an alternative approach of utilizing direct responses to a survey on the earnings test, which is unique in our dataset compiled by the Japanese Government, and provide new evidence on the sensitivity of the labor supply decision of workers aged between 60 and 64 with respect to the earnings test. We take advantage of micro-level data from the nationwide survey on employment of the elderly to examine the change in the labor supply effect

for those aged 60–64 in the 1996 and 2000 surveys.

Our analysis provides several important findings. First, half of those workers are discouraged from working or reduce their working hours. This is also the case even after correcting for the observable attributes among individuals who reported affected and unaffected. Second, the revision of the test rules in 1995 did not alter the labor supply of the elderly.

Our empirical findings show that the labor supply effect of the social security earnings test may have been underestimated because of measurement errors or labor market rigidities. In this study, we argue that a direct response to the question on the labor supply effect is an alternative to complement the traditional methodology to examine the labor supply effect of the earnings test. Further study should examine the effect of the social security earnings test considering other important factors influencing the labor supply decision such as health status, family relationships, and labor-leisure choices as well as the retirement decision of the elderly.

References

- Abe, Yukiko (1998). “1980–1990 Nendai no Dansei Koureisha no Roudou Kyokyu to Zaishoku Rorei Nenkin Seido” (Labor Supply of the Elderly Male and the Earnings-Tested Pension in the 1980s and 1990s) [in Japanese], *JCER Economic Journal (Nihon Keizai Kenkyu)*, vol. 36, pp. 50–82.
- Baker, Michael and Dwayne Benjamin. 1999. “How Do Retirement Tests Affect the Labour Supply of Older Men?” *Journal of Public Economics* 71(1):27–52.
- Borjas, George J. (2005). *Labor Economics* (Third Edition). Boston: McGraw-Hill Irwin Press.
- Burtless, Gary, and Robert A. Moffitt (1985). “The Joint Choice of Retirement Age and Post-Retirement Hours of Work,” *Journal of Labor Economics*, vol. 3, pp. 209–236.
- DiNardo, John (2002). “Propensity Score Reweighting and Changes in Wage Distributions,” University of Michigan, mimeograph.
- DiNardo, John, Nicole Fortin and Thomas Lemieux (1996). “Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach,” *Econometrica*, vol. 65 no. 5, pp. 100–1046.
- Disney, Richard and Sarah Smith (2002). “The Labour Supply Effect of the Abolition

- of the Earnings Rule for Older Workers in the United Kingdom,” *Economic Journal*, vol. 112, issue 478, pp. C136–C152.
- Friedberg, Leona (2000). “The Labor Supply Effects of the Social Security Earnings Test,” *Review of Economics and Statistics*, vol. 82, no. 1, pp. 48–63.
- Gruber, Jonathan and Peter Orszag (2003). “Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?” *National Tax Journal*, vol. 56, no. 4, pp. 755–773.
- Gustman, Alan L., and Thomas L. Steinmeier (1985). “The 1983 Social Security Reforms and Labor Supply Adjustments of Older Individuals in the Long Run,” *Journal of Labor Economics*, vol. 3, pp. 237–253.
- Haider, Steven and David Loughran (2008). “The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data,” *Journal of Human Resources*, vol. 43, no. 1, pp. 57–87.
- Iwamoto, Yasushi (2000). “Zaishoku Rorei Nenkin Seido to Koureisha no Shugyo Kodo” (The Social Security Earnings Test and Labor Supply of the Elderly) [In Japanese], *Quarterly of Social Security Research (Kikan Shakai Hoshō Kenkyū)*, vol. 35, no. 4, pp. 364–376.
- Lemieux, Thomas (2002). “Decomposing Changes in Wage Distributions: A Unified

Approach,” *Canadian Journal of Economics*, vol. 35, no. 4, pp. 646–688.

Ogawa, Hiroshi (1998). “Nenkin ga Koureisha no Shugyo Kodo ni Ataeru Eikyo ni tsuite” (On the Impact of the Public Pension on Labor Supply of the Elderly) [in Japanese], *Keizai Kenkyu (Economic Review)*, vol. 49, no. 3, pp. 245–258.

Shimizutani, Satoshi and Takashi Oshio (2008). “The Labor Supply Effect of Social Security Earnings Test Revisited: New Evidence from its Elimination and Revival in Japan,” Project on Intergenerational Equity and Center for Intergenerational Studies (CIS), Institute of Economic Research, Hitotsubashi University, Discussion paper, no. 410.

Shimizutani, Satoshi and Izumi Yokoyama (2009), “Has Japan’s Long-Term Employment Practice Survived? Developments since the 1990s,” *Industrial and Labor Relations Review*, vol. 62, no. 3, April 2009, pp. 311–324.

Table 1 Direct response to the earnings test in the 2000 survey

	No work group	Restricted group	No effect group	MAI beneficiaries
Wage income	1.14	13.58	27.22	11.82
(ten thousand yen)	(5.83)	(12.73)	(26.79)	(19.33)
Nonwage income	14.50	15.72	8.37	20.71
(ten thousand yen)	(10.91)	(9.62)	(10.476)	(6.85)
Wage + pension income	11.59	25.95	34.23	31.75
(ten thousand yen)	(9.48)	(14.94)	(26.40)	(19.23)
Working hours per day	0.45	6.65	7.53	3.87
	(1.80)	(2.07)	(2.09)	(3.76)
Working days per week	0.30	4.29	5.17	2.58
	(1.21)	(1.39)	(1.24)	(2.53)
Monthly working hours	2.14	28.790	39.86	18.40
	(9.20)	(14.03)	(15.46)	(20.23)
Number of observations	576	322	892	321

Note: Figures in parentheses are the standard deviation.

Table 2 Summary statistics in the 2000 survey

	Those affected by the earnings test		Those not affected by the earnings test	
	mean	S.D.	mean	S.D.
The affected*	1	0	0	0
Wage income	5.602	10.745	27.217	26.787
Nonwage income	14.939	10.477	8.365	10.476
Current EPI benefits	10.057	8.598	3.070	5.854
Current earnings-tested EPI benefits	0.728	2.890	1.734	3.688
Full benefits without the earnings test	1.079	4.302	3.942	7.439
Male *	0.609	0.488	0.723	0.448
Age 60 (dummy variable)*	0.187	0.390	0.221	0.415
Age 61 (dummy variable)*	0.185	0.388	0.192	0.394
Age 62 (dummy variable)*	0.212	0.409	0.203	0.402
Age 63 (dummy variable)*	0.232	0.422	0.215	0.411
Age 64 (dummy variable)*	0.185	0.388	0.169	0.375
Educational attainment (junior high school)	0.411	0.492	0.352	0.478
Educational attainment (senior high school or two-year college)*	0.497	0.500	0.474	0.500
Educational attainment (university)*	0.089	0.285	0.170	0.376
Health status (healthy)*	0.709	0.454	0.855	0.352
Health status (not healthy)*	0.188	0.391	0.121	0.326
Health status (sick)*	0.102	0.303	0.024	0.152
Physical status (possible to work on a full-time basis)	0.315	0.465	0.674	0.469
Physical status (possible to work depending on work conditions)	0.513	0.500	0.320	0.467
Physical status (impossible to work)	0.171	0.377	0.007	0.082
Family size	2.843	1.382	2.889	1.433
Number of observations	898		892	

Note: * refers to a dummy variable. The summary statistics for each prefecture dummy (47 dummies) are omitted.

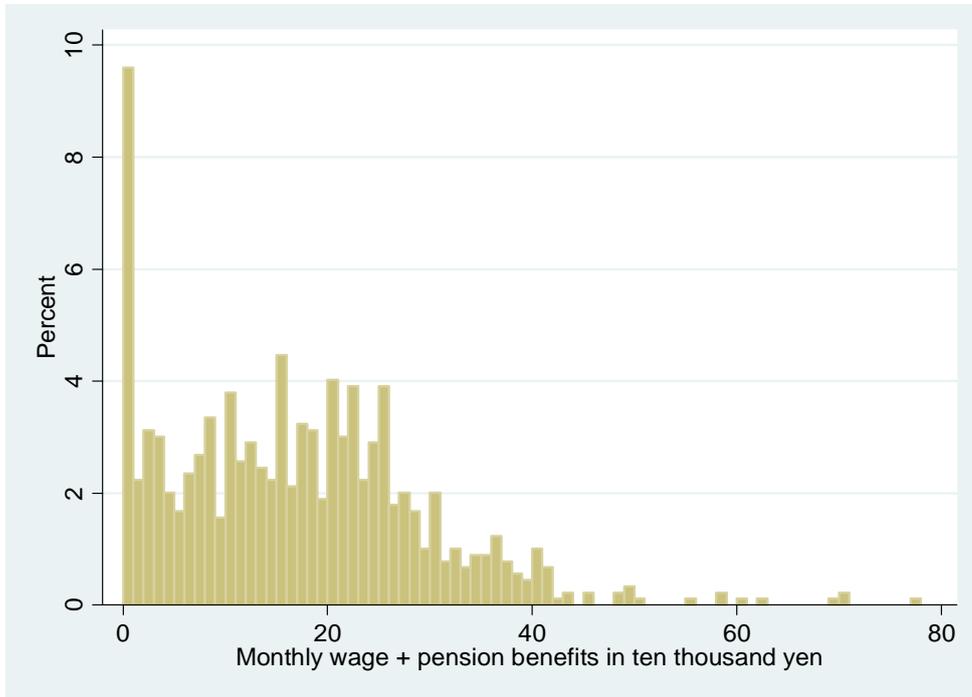
Table 3 Summary statistics in the 1996 survey

	Those affected positively		Those not affected at all	
	mean	S.D.	mean	S.D.
The affected*	1	0	0	0
Wage income	18.882	8.146	18.838	15.348
Nonwage income	11.000	6.831	9.618	7.565
Current earnings-tested EPI benefits	8.559	5.304	7.544	5.371
Full benefits without the earnings test	14.629	7.693	13.953	9.919
Male *	0.941	0.239	0.724	0.448
Age 60 (dummy variable)*	0.265	0.448	0.171	0.377
Age 61 (dummy variable)*	0.353	0.485	0.215	0.412
Age 62 (dummy variable)*	0.176	0.387	0.215	0.412
Age 63 (dummy variable)*	0.088	0.288	0.202	0.402
Age 64 (dummy variable)*	0.118	0.327	0.197	0.399
Health status (healthy)*	0.853	0.359	0.842	0.365
Health status (not healthy)*	0.147	0.359	0.145	0.353
Health status (sick)*	0.000	0.000	0.013	0.114
Physical status (possible to work on a full-time basis)	0.765	0.431	0.561	0.497
Physical status (possible to work depending on work conditions)	0.235	0.431	0.430	0.496
Physical status (impossible to work)	0.000	0.000	0.009	0.093
Family size	2.735	1.463	3.181	1.604
Number of observations	34		228	

Note: * refers to a dummy variable. The summary statistics for each prefecture dummy (47 dummies) are omitted from this table.

Figure 1 Distribution of wage plus pension benefits

(A) Those who are affected by the earnings test



(B) Those who are not affected by the earnings test

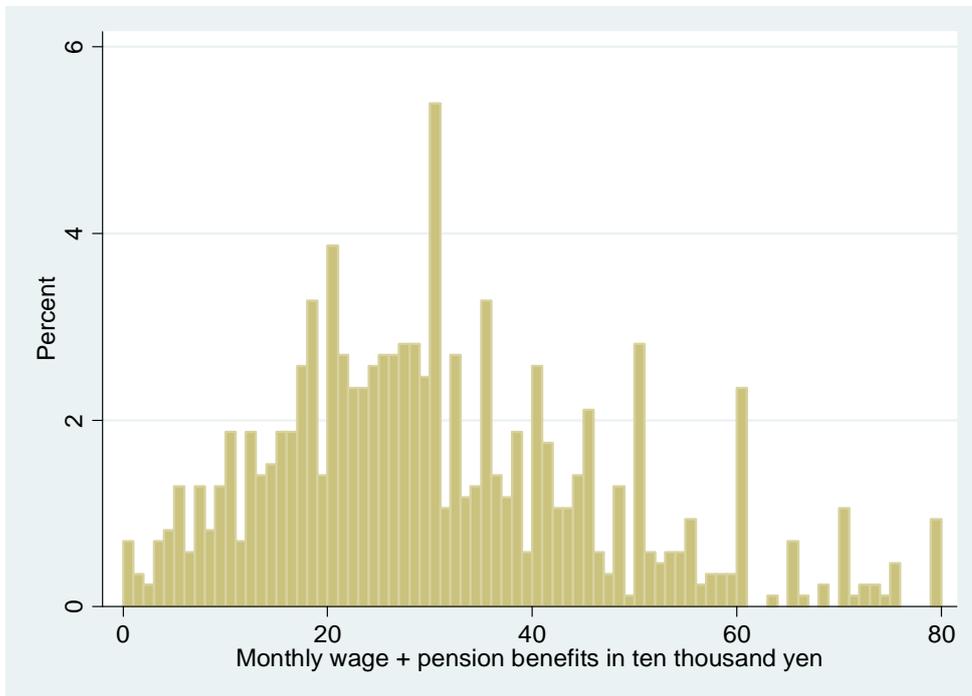
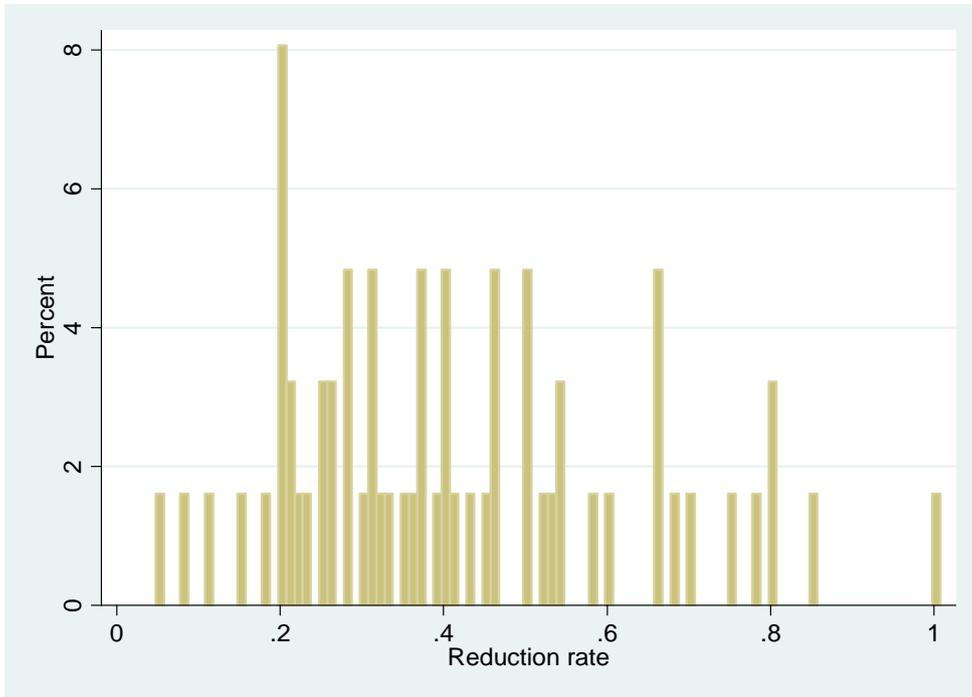


Figure 2 Distribution of reduction rates by the earnings test

(A) Those who are affected by the earnings test



(B) Those who are not affected by the earnings test

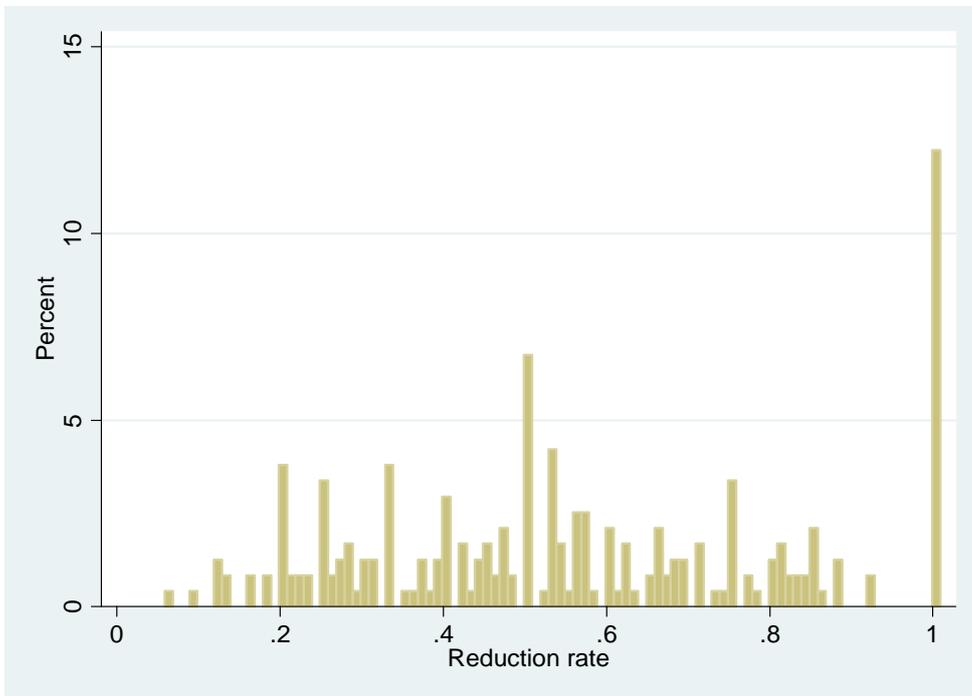


Figure 3 Decomposition analysis of wage and pension benefits

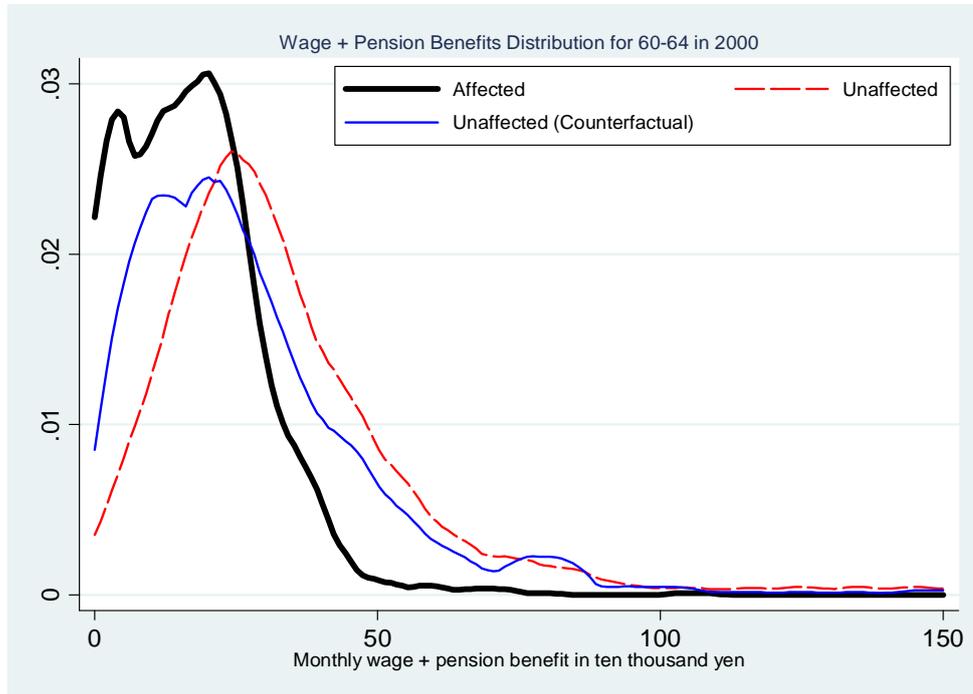


Figure 4 Decomposition analysis of wage and pension benefits excluding non workers

